

## $\alpha_{IIb}\beta_3$ 同源模拟和环形RGD药物小分子设计

罗明艳<sup>1,2</sup>, 陈梅宗<sup>1,2</sup>, 李任植<sup>1,2,3</sup>

1 南开大学物理科学学院, 天津 300071

2 南开大学泰达应用物理学院, 天津 300457

3 南开大学生物活性材料教育部重点实验室, 天津 300071

**摘要:** 血小板表面的整合蛋白 $\alpha_{IIb}\beta_3$ 对血栓的形成具有很重要的调控作用, $\alpha_{IIb}\beta_3$ 与纤维蛋白原上的RGD特征序列结合,促进血小板凝集进而形成血栓。这样可以在理论上设计一个环形RGD小分子(RGD-c)与 $\alpha_{IIb}\beta_3$ 蛋白结合,阻断其与纤维蛋白原的结合位点,进而阻断血栓形成。以糖蛋白 $\alpha_v\beta_3$ (pdb 代码 1jv2)作为模板,利用modeller8v2软件进行同源模拟得到 $\alpha_{IIb}\beta_3$ 三维模型。而小分子则以RGD序列为主,两边各加一个氨基酸X、Y, X和Y之间用一个二硫键相连,通过改变X和Y,重复优化过程则可得到不同的小分子模型,将其与 $\alpha_{IIb}\beta_3$ 进行分子对接,可以找出比较理想的XY组合,对治疗血栓的药物设计具有重要的理论指导作用

**关键词:**  $\alpha_{IIb}\beta_3$ , 同源模拟, 对接, RGD-c, MM

## $\alpha_{IIb}\beta_3$ Modeling Simulation and Oesign of the Cyclic RGD

Mingyan Luo<sup>1</sup>, Meizong Chen<sup>1</sup> and Imshik Lee<sup>1,2,3</sup>

1 College of Physics, Nankai University, Tianjin 300071, China

2 TEDA Applied Physics School, Nankai University, Tianjin 300457, China

3 Key Laboratory of Bioactive Materials of Ministry of Education, Nankai University, Tianjin 300071, China

**Astract:** Integrin  $\alpha_{IIb}\beta_3$  of the platelet surfaces regulates the thrombosis formation.  $\alpha_{IIb}\beta_3$  binds to the RGD sequence (Arg-Gly-Asp) of fibrinogen, promotes the platelet aggregation and finally leads to the thrombus. We obtained the three-dimensional molecular structure of  $\alpha_{IIb}\beta_3$  using homology-modeling (modeller8v2 software), with integrin  $\alpha_v\beta_3$  (pdb code 1JV2) as the template. Accordingly, a cyclic RGD(RGD-c) peptide was designed to bind  $\alpha_{IIb}\beta_3$  as an antagonist and to block the formation of thrombus. We added two amino acids X, Y to both sides of RGD-c. X and Y could bind to each other by disulfide bond that finally made RGD-c a cyclic peptide. The optimum structure of RGD-c was obtained from the energetic optimization processes. All amino acids were placed at the X and Y to conduct Molecular Docking to the integrin  $\alpha_{IIb}\beta_3$ . We got the optimum structure of RGD-c by energetic optimization and the antagonistic combination analysis. The results might provide an insight into designing and screening integrin  $\alpha_{IIb}\beta_3$  antagonists.

**Keywords:**  $\alpha_{IIb}\beta_3$ , homology modeling, dock, RGD-c, MM

$\alpha_{IIb}\beta_3$ 是血小板上的一种重要的受体蛋白<sup>[1]</sup>,由两个亚基组成,可以与纤维蛋白原上的RGD特征序

列结合,使血小板凝集,其活性的异常表达是血栓形成的重要因素。在过去的20年里,人们对于 $\alpha_{IIb}\beta_3$ 的

Received: June 13, 2007; Accepted: August 6, 2007

Supported by: the Doctoral Fund of the Ministry of Education (No. 20050055037).

Corresponding author: Imshik Lee. Tel: +86-22-23494411; E-mail: ilee@nankai.edu.cn

国家自然科学基金资助 (No. 20050055037).

结构已经有了很多的认识,特别是它的同源蛋白 $\alpha_v\beta_3$ 的结构<sup>[2]</sup>已经通过X射线衍射的方式得到,可以从其中得出很多关于前者的信息,比如用同源模拟<sup>[3]</sup>的方法就可以得到它的高精度模型。然后根据蛋白受体和配体竞争结合<sup>[4]</sup>的理论,设计一些与纤维蛋白活性部位相似但氨基酸数目少,结构比较稳定的小分子<sup>[5]</sup>,或者其他非肽类小分子作为血栓形成信号通路中某一活性蛋白的配体拮抗剂,可以有效防止血栓的形成,而且副作用小。比如蛋白激酶C(PKC)能够活化 $\alpha_{IIb}\beta_3$ ,使之开放与纤维蛋白原的结合位点,使用PKC的阻断剂可以明显降低血小板凝集的水平,但是这种方式会大量阻断PKC家族的其它作用,影响其正常生理功能的表达,因此我们可以选择另一种方式,用 $\alpha_{IIb}\beta_3$ 自身的阻断剂来阻断纤维蛋白原与之的结合。目前,已经有多种该蛋白的拮抗剂类药物存在,比如盐酸替罗非班,阿昔单抗等,实用量很少即可有效延迟血栓形成时间,减轻血栓重量。血小板上受体蛋白拮抗剂类抗凝药物基本上已经进入了实用阶段。

如上所说, $\alpha_{IIb}\beta_3$ 作为凝血过程中的重要一员,可以用药物阻断其活化过程,但是目前的阻断剂基本上都是非肽类可逆小分子,虽然能阻断大部分活化信号,但其作用特异性不够,容易产生副作用。相比之下,肽类小分子可以阻断特定受体,结合特异性强,属于指哪打哪型药物,不会对其他信号或是活化通路发生作用,而且组成也不太复杂,往往由几十个甚至几个氨基酸构成,可以有效的规避风险。在血小板凝集过程中,能与 $\alpha_{IIb}\beta_3$ 特异性结合的特征序列有两种,一种是通过KQAGDV特征序列与 $\alpha_{IIb}\beta_3$ 蛋白上的特异性结合位点相互作用,另一种就是前面提到的RGD特征序列<sup>[6]</sup>。已经有学者用RGD-葡激酶突变体来抗血小板聚集<sup>[7]</sup>。本文将系统建立X-RGD-Y模型<sup>[8]</sup>,通过分子模拟对接的方式,了解其与 $\alpha_v\beta_3$ 结合的稳定度和特异性,分析其活化,结合位点和结合原理,找出与 $\alpha_v\beta_3$ 结合的最佳氨基酸组合。为血栓类抗凝肽类小分子药物的实验室筛选和配制提供理论支持,减少实验室操作时间和经费投入。

## 1 方法

### 1.1 $\alpha_{IIb}\beta_3$ 的同源模拟

#### 1.1.1 $\alpha$ 亚基中活性区域的选取

$\alpha_v\beta_3$ 和 $\alpha_{IIb}\beta_3$ 蛋白均由 2 个相关亚基组成,其中

的 $\beta$ 亚基是基本相同的。首先考虑 $\alpha$ 亚基,前者是 1039 个氨基酸,后者是 927 个氨基酸(不包括胞内部分)。 $\alpha_v$ <sup>[8]</sup>由 4 个结构域组成,分别是thigh, calf-1, calf-2 以及 $\beta$ -propeller。其中 $\beta$ -propeller是整个 $\alpha$ 亚基中主要的活性区域。从N末端开始,每隔 60 个左右氨基酸重复 1 次,一共 7 组,每组由 1 个 4 次反向折叠的 $\beta$ 片层组成,呈放射状的构成 1 个 7 叶螺旋桨的结构。其中每一组的里面片层组成该区域的内部结构,中间两个则逐渐向外扩展,而最后 1 个作为上一组的类似重复,与下一组相邻。另外整个 $\beta$ -propeller还包含 4 个钙离子结合位点。

用modeller程序进行序列比对,其序列相似性为 37%,如果除去 $\alpha_{2b}$ 对应的胞内部分,相似性在 40%以上。我们用模拟软件modeller8v2 对目标蛋白质进行模拟,得到其模拟结构。也分为 4 个区域,其中包括 $\beta$ -propeller,另外 3 个区则分别类似于 $\alpha_v$ 的 thigh区, calf-1 和calf-2 区。

在该亚基上,有一定区域的跨膜部分,由于模版在同样位置的结构尚不清楚,模拟结果中有相当部分是空白的,仅仅是自身优化的结果,在后面的对接试验中被认为是人为操作的结果不予接受,所以将 $\alpha$ 亚基的唯一功能区域剥离,重点对该部分进行模拟优化,这就是 $\beta$ -propeller<sup>[10]</sup>。如图 1 所示。

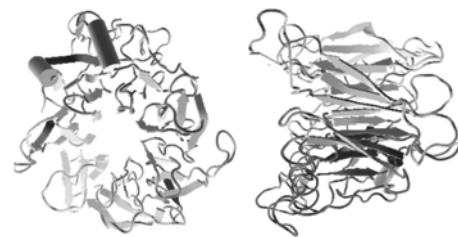


图 1  $\beta$ -propeller 上视图和侧视图

Fig.1 The look-up and look-down structure of  $\beta$ -propeller

运用 arguslab 软件的几何结构优化命令,UFF 力场下进行最速下降优化,迭代 1000 次;BFGS 优化,迭代 1000 次,非键作用力设定 8-10 埃。Spdbv 软件执行能量最小化操作一次。优化前后 $\beta$ 片层无变化,回折区域结构变化明显,出现明显的 $\alpha$ 螺旋结构。优化完成后,得到的 $\beta$ -propeller 与模版结构基本相同,也是 7 叶螺旋桨结构, $\alpha$ 螺旋结构在第三组叶片中间。前六组叶片包括 4 个 $\beta$ 片层,由内到外顺次展开,最后一组只有 3 片,并与下一区域连接。

### 1.1.2 $\beta$ 亚基中活性区域的选取

$\beta_3$ 亚基分为 $\beta A$ , EGF-3, EGF-4,  $\beta TD$ 和hybrid这5个区域<sup>[11]</sup>。其中 $\beta A$ 与 $\alpha$ 亚基的 $\beta$ -propeller区相连,两者组成整个蛋白中活性最强的部分, $\alpha_{IIb}\beta_3$ 几乎所有的活性位点都在这两个区中,其中就包括金属离子依赖性支持位点(MIDAS—Asp119, Ser121, Ser123, Glu220 以及 Asp251)以及一个金属离子结合位点。由于模板和目标蛋白的 $\beta_3$ 亚基序列相似度高达98%,在已知的功能区域 $\beta A$ 区,两者是完全一样的,这里直接使用模板对应区域的结构(如图2)。

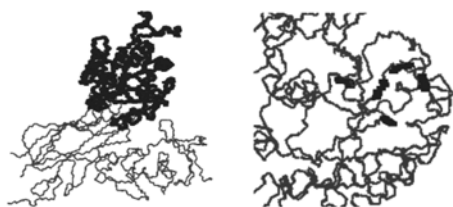


图2 亚基结构(蓝色为 $\beta A$ )和 $\beta A$ (蓝色为 MIDAS 区)

Fig.2 The structure of subunit  $\beta_3$  and  $\beta A$

The blue stead for  $\beta A$  in the left structure and MIDAS in the right one

由于本论文的目的是检验 RGD 序列两边 X, Y 对结合特异性和稳定度的影响,后面的分子对接考虑的也是包含可能的结合部位的区域,因此对于整体的精确度没有多高的要求。在模拟过程中,我们挑出两个亚基中需要做分子对接的最重要的两个区域进行了重点模拟和优化。整体结构只是在各个区域模拟完成之后作了一个比较简单的整合,而且对应模板未知部分的序列被删去了。

## 1.2 环形小分子 X-RGD-Y 的设计

### 1.2.1 主链结构模拟

所有小分子主链由Modeller8v2完成,操作方法与 $\alpha_{IIb}\beta_3$ 相同。对于5氨基酸短肽链来说,Modeller8v2的模拟过程显得非常粗糙,但是本题中这一步的目的只是将5个氨基酸按顺序连接起来,其具体结构会在后面进行修正。

### 1.2.2 添加二硫键

利用arguslab的add atom功能,先将C末端的碳原子上-OH和N末端氮原子上一个-H去掉,然后在合适位置添加两个硫原子,将C末端碳原子,S-S,N末端氮原子依次连接即可。

### 1.2.3 结构优化

每个小分子优化过程均相同,分子力学优化,

全力场最速下降1000次迭代,BFGS优化1000次迭代。能量最小化操作一次。由于分子小,而且认为添加了2个硫原子和3个单键,优化过程非常明显,从线形结构到环形结构各个参数均有较大幅度的变化,如图3。

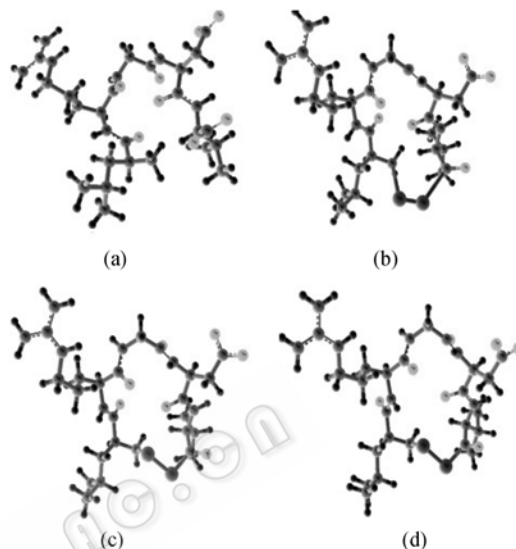


图3 初始结构(a)添加硫原子(蓝色)(b)最速下降优化(c)BFGS优化(d)

Fig.3 initial structure(a), adding sulfur atoms(b), the optimal by the steepest descent method(c), and BFGS(d)

## 1.3 分子对接

全部的分子对接试验均在Gramm服务器http://www.bioinformatics.ku.edu完成,返回对接完成后的pdb文件。该服务器采用高精度对接模拟,全局搜索所有可能的结合方式,保留并返回十种最有利的对接结果。

## 2 结果与讨论

### 2.1 $\alpha_{IIb}$ 亚基与配体小分子的对接

$\alpha_{IIb}$ 与配体的结合主要集中在 $\beta$ -propeller外侧以及各组刀片的 $\beta$ 回折部分,到目前为止,其与纤维蛋白原结合机制未明,尚未有证据表明其与RGD特征序列有特异性结合作用。本题以其与RGD小分子做分子对接模拟,仅仅是作为一个试验性的参考,观察其上是否存在RGD结合位点,找出可能性的结合区域。

在实验中,把20种氨基酸分为疏水和亲水两组,结果证明,只有亲水性氨基酸能与该区域活性部位结合。而选用疏水性氨基酸的配体小分子绝大部分

都被包埋到 $\beta$ -propeller 内侧, 仅与 $\beta$ 片层之间有范德华力的作用因此第一阶段即被排除。后期试验全部选用亲水性氨基酸。亲水性氨基酸一共 12 个, 排列组合 132 组, 与受体分子做分子对接 132 次, 返回结果 pdb 文件 132 组 1320 个。由于目前该区域的活性位点未明, 只能保留所有与该区域外侧结合的结果作统计分析(如图 4)。

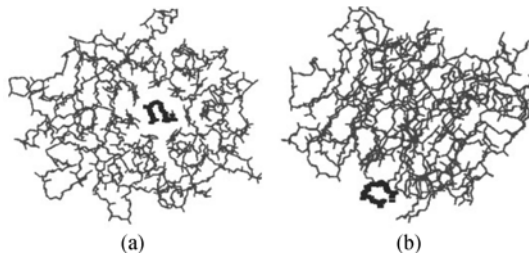


图 4 RGD 与 $\beta$ -propeller 结合图

Fig.4 the binding structure RGD and  $\beta$ -propeller

(a) the ligand binding with the inner of  $\beta$ -propeller, as there are 7 group of brand in inner place, no special binding site (b) the binding sites located in the outer of brand. we statistic this kind of binding

我们对所有结合位点进行统计, 在 124Ala, 129Trp, 200Phe, 205Thr, 206Gln, 242Leu-244His, 266Tyr, 329Gly-331Gly, 387Ser, 393Gly, 401Asn, 448Phe, 453Ala, 454Val 这一系列残基上, 出现的结合次数最多, 在所有结果中出现的百分比在 10%~20%之间, 而其他残基的出现概率在 10%以下。此外, 在上面的统计中, 除了 242Leu-244His 和 329Gly-331Gly 外, 其他的残基均为折叠两侧残基, 其单独作为结合位点的可能性较小, 故认为, 如果该区域存在 RGD 特异性结合位点, 该位点很可能在 242Leu-244His 和 329Gly-331Gly 附近。

目前该亚基的结构均基于程序模拟和假定的基础上, 本文只能保证在现有的条件上达到最高的精度, 其细微结构还需要更多的真实信息加以佐证。在后半段的实验过程中, 可以在重点分析 $\beta_3$ 亚基结合方式的同时, 注意该亚基上是否有特异性结合。

## 2.2 $\beta_3$ 亚基与配体小分子的对接

该部分的对接模拟, 20 种氨基酸, 一共是 400 种组合。由于 $\beta_3$ 亚基上的 RGD 特征性结合位点已经清楚, 在 4000 个结果中进行分析, 发现对接结果相当稳定, 绝大部分结合在 $\beta$ A 区域与下一区的连接处的被我们第一位排除, 剩下的大概占总结果的 10%, 主要结合部位就是在已知的活性位点上。在这部分

保留的结果中, 作了详细的观察和分析。因为采用的小分子只有五个氨基酸, 所以我们逐个检查每个氨基酸对应的结合位置, 与已知结果相对应, 排除非正常结合, 包括结合错位, 小分子结合位置偏移过大等(如图 5)。

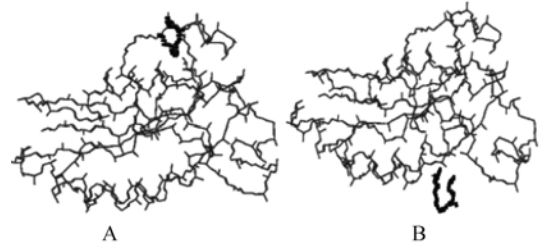


图 5 RGD 与 $\beta$ A 的结合图

Fig.5 The binding structure RGD and  $\beta$ A

A: no binding sites in the connection and the next domain in the theory, so this kind of binding is deleted; B: the right sites but the wrong amino acid (the open place towards up, the binding amino acid is RGD)

在对接结果中, 检验对接效果的一个重要依据就是结合自由能非键能和受体间氢键数量。所以我们利用 SPDBV 软件在 GROMOS96 下计算分子间结合非键能<sup>[12]</sup>(氢键能和范德华能), 得到下面的非键能图 6。

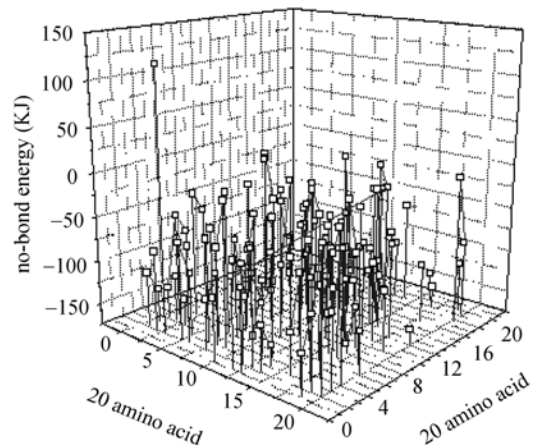


图 6 在已知结合位点有 RGD 小分子分布图

Fig. 6 the distribute of RGD in the binding site

The arrangement of 20 amino acid by hydrophile to hydrophile  
AVLIPWFMGSTCQYDRKHEN

$\beta_3$ 亚基没有模拟过程, 试验主要是对接, 根据上图数据分析可以得到以下几点:

(1) 包含 RGD 序列的 5 氨基酸小分子可以与其特性结合, 结合位置与已知结果相符。具体氨基酸对应的受体上结合残基与资料结果有出入, 但差别

不大。可以确定该亚基上 RGD 特异性结合位点在 125Lys-(DDL)-129Trp, 132Gln, 208Lys, 210Gln-211 Ser-212 Val 段。

(2) RGD 两侧选用不同的氨基酸, 对特异性结合有很大影响。绝大部分组合不能结合在正确位置, 占总数的 70%。

(3) 从统计结果看, 选用亲水性氨基酸的效果要好于亲疏结合的, 疏水的最次。其中非键结合能较高的小分子中, 代正电荷的亲水精氨酸、赖氨酸, 代负电荷的亲水谷氨酸、天冬氨酸, 还有带有苯环的苯丙氨酸、Y 搭配组成的小分子具有较高的非键结合能。

(4) 综合起来, 5aa 小分子其两侧氨基酸不足以突出 RGD 中心对两边的的选择作用, 而且分子太小, 其可以结合的部位增加, 降低了特异结合可能性; 根据受体结合部位大致 10 个的结合氨基酸构成区域来看, 配体小分子氨基酸数也应该在 10 个左右, 这样在保证小分子活性的同时, 可以防止由于肽链的加长导致受体与其非 RGD 序列结合的可能性增加。

### 3 展望

综上所述, 本文主要研究了  $\alpha_{IIb}\beta_3$  蛋白的结构以及其与 RGD 小分子药物的对接模拟。通过同源模拟的方法得到模拟结构, 计算模拟其与不同小分子特异性结合的稳定性及可靠性, 研究得到了一些有用信息, 对大量实验结果进行了有效处理, 由于试验结果数量庞大, 处理结果时不得不采用统计分析的方法排除掉大部分, 可能会漏掉有用信息。虽然由于条件和时间的限制还存在很多不足, 但对于用实验来验证受体的结合位点方面可以减小很多工作量, 可以直接针对模拟的结果来进行试验。本题中采用的小分子均为 5 氨基酸结构, 根据肽链功能与其长度的正比关系以及结合复杂程度与其长度的反比关系, 下一步需要增加氨基酸数目, 测试不同数量氨基酸链与受体结合的效果, 找出比较合适的肽链长度, 根据试验结果配药, 实验室条件下测试药物分

子与凝血时间以及血栓大小的关系, 为研制新药做出一定的理论基础贡献。

### REFERENCES

- [1] Seymour JL, Henzel WJ, Nevins B. Decorsin, a potent glycoprotein IIb/IIIa antagonist and platelet aggregation inhibitor from the leech *Macrobdella decora*. *J Biol Chem* 1990; **265**: 10143-7.
- [2] Xiong JP, Stehle T, Diefenbach B, *et al.* Crystal Structure of the extracellular segment of integrin  $\alpha V/\beta 3$ . *Science* 2001, **294**: 339-46.
- [3] Greer J. Comparative Modeling of Homologous Proteins. *Meth. Enzymol.* 1991, **202**: 239-252.
- [4] Yang J. Interaction of human fibrinogen receptor (GPIIb-IIIa) with decorsin. *Acta Pharmacol Sin* 2004, **25**: 1096-1104.
- [5] Blundell, TL, Sibanda, BL, Sternberg, MJE. Knowledge-Based Prediction of Protein Structures and the Design of Novel Molecules. *Nature* 1987, **326**: 347-352.
- [6] Topol EJ, Byzove TV, Plow EF. Platelet GPIIb/IIIa blockers. *Lancet* 1999, **353**: 227-31.
- [7] Ning BA, Ma R, Zheng YL, *et al.* Preparation and activity analysis of RGD-mSAK(K130T, K135R). *Chinese Journal of Biotechnology*, 2005(21), **3**: 456-460.  
宁保安, 马茹, 郑玉玲, 等. RGD-葡激酶突变体(K130T, K135R)的制备与活性分析. *生物工程学报*, 2005, **3**(21): 456-460.
- [8] Polverino de Laureto P, Scaramella E, De Filippis V. Chemical synthesis and structural characterization of the RGD-protein decorsin: a potent inhibitor of platelet aggregation. *Protein Sci* 1998, **7**: 433-44.
- [9] Xiong JP, Stehle T, Diefenbach B, *et al.* Crystal structure of the extracellular segment of integrin  $\alpha V\beta 3$ . *Science*, 2001: 339-345.
- [10] Vilmos Fulop, David TJ.  $\beta$ -propellers: structural rigidity and functional diversity. *Current Opinion in Structural Biology*, 1999, **9**: 715-721.
- [11] Juan J. Calvete platelet integrin GPIIb/IIIa: structure function correlation an update and lessons from other integrins. *Instituto de Biomedicina de Valencia, C.S.I.C.*
- [12] Morris GM, Goodsell DS, Halliday RS, *et al.* Automated docking using a Lamarckian genetic algorithm and empirical binding free energy function. *J Computat Chem*, 1998, **19**: 1639-1662.